

ABSTRACT

Common diseases have several symptoms or initial indications before a disease occurs that can threaten health conditions. Some people still take it for granted and are reluctant to have their health checked by a doctor because of various conditions. The lack of distribution of doctors in several regions in Indonesia also affects health services to the community. This research aims to develop a web-based system for diagnosing common diseases based on symptoms using the Decision Tree classification method. This research took the research object at the Jemursari Islamic Hospital, Surabaya, using a dataset in the form of 9934 patient medical records. The attribute datasets used in this system include complaints, rps, and ratings. This system was built using the Decision Tree model as a data mining algorithm, with Python, HTML, CSS, and JS as programming languages. The general results obtained from the disease diagnosis system based on web-based complaints with a comparison of the C4.5 and CART methods were able to produce an accuracy of 96.045197740113% and 96.61016949152542%. Each algorithm performed well in classifying diseases based on patient symptoms, with the CART algorithm showing slightly better accuracy. Then the performance of the C4.5 model was evaluated using cross-validation, which showed that the model did not overfit significantly. The error per fold is stable, with an average error value of 0.0569, indicating that the model is not too sensitive to changes in data and can work well on new data. The training error value of 0.0242 is lower than the testing error value of 0.0339. However, the difference is not significant, being only 0.0097, which indicates the model's ability to generalize well to new data.

Keywords— Data Mining, Decision Tree, Disease Diagnosis System, Cross Validation, Confusion Matrix